



SEQUENCE LISTING

<110> Barany, Francis
Cao, Weiguo
Tong, Jie

<120> HIGH FIDELITY THERMOSTABLE LIGASE AND USES THEREOF

<130> 19603/2615

<140> 09/830,502

<141> 1999-10-29

<150> 60/106,461

<151> 1998-10-30

<150> PCT/US99/25437

<151> 1999-10-29

<160> 24

<170> PatentIn Ver. 2.1

<210> 1

<211> 674

<212> PRT

<213> Thermus sp.

<400> 1

Met Thr Leu Glu Glu Ala Arg Arg Arg Val Asn Glu Leu Arg Asp Leu
1 5 10 15

Ile Arg Tyr His Asn Tyr Leu Tyr Tyr Val Leu Asp Ala Pro Glu Ile
20 25 30

Ser Asp Ala Glu Tyr Asp Arg Leu Leu Arg Glu Leu Lys Glu Leu Glu
35 40 45

Glu Arg Phe Pro Glu Leu Lys Ser Pro Asp Ser Pro Thr Glu Gln Val
50 55 60

Gly Ala Arg Pro Leu Glu Ala Thr Phe Arg Pro Val Arg His Pro Thr
65 70 75 80

Arg Met Tyr Ser Leu Asp Asn Ala Phe Ser Leu Asp Glu Val Arg Ala
85 90 95

Phe Glu Glu Arg Ile Glu Arg Ala Leu Gly Arg Lys Gly Pro Phe Leu

100	105	110
Tyr Thr Val Glu Arg Lys Val Asp Gly Leu Ser Val Asn Leu Tyr Tyr		
115	120	125
Glu Glu Gly Ile Leu Val Phe Gly Ala Thr Arg Gly Asp Gly Glu Thr		
130	135	140
Gly Glu Glu Val Thr Gln Asn Leu Leu Thr Ile Pro Thr Ile Pro Arg		
145	150	155
Arg Leu Thr Gly Val Pro Asp Arg Leu Glu Val Arg Gly Glu Val Tyr		
165	170	175
Met Pro Ile Glu Ala Phe Leu Arg Leu Asn Gln Glu Leu Glu Ala		
180	185	190
Gly Glu Arg Ile Phe Lys Asn Pro Arg Asn Ala Ala Ala Gly Ser Leu		
195	200	205
Arg Gln Lys Asp Pro Arg Val Thr Ala Arg Arg Gly Leu Arg Ala Thr		
210	215	220
Phe Tyr Ala Leu Gly Leu Gly Leu Glu Thr Gly Leu Lys Ser Gln		
225	230	235
His Asp Leu Leu Leu Trp Leu Arg Glu Arg Gly Phe Pro Val Glu His		
245	250	255
Gly Phe Thr Arg Ala Leu Gly Ala Glu Gly Val Glu Glu Val Tyr Gln		
260	265	270
Ala Trp Leu Lys Glu Arg Arg Lys Leu Pro Phe Glu Ala Asp Gly Val		
275	280	285
Val Val Lys Leu Asp Asp Leu Ala Leu Trp Arg Glu Leu Gly Tyr Thr		
290	295	300
Ala Arg Thr Pro Arg Phe Ala Leu Ala Tyr Lys Phe Pro Ala Glu Glu		
305	310	315
Lys Glu Thr Arg Leu Leu Ser Val Ala Phe Gln Val Gly Arg Thr Gly		
325	330	335
Arg Ile Thr Pro Val Gly Val Leu Glu Pro Val Phe Ile Glu Gly Ser		
340	345	350
Glu Val Ser Arg Val Thr Leu His Asn Glu Ser Phe Ile Glu Glu Leu		

355	360	365
Asp Val Arg Ile Gly Asp Trp Val Leu Val His Lys Ala Gly Gly Val		
370	375	380
Ile Pro Glu Val Leu Arg Val Leu Lys Glu Arg Arg Thr Gly Glu Glu		
385	390	395
Lys Pro Ile Ile Trp Pro Glu Asn Cys Pro Glu Cys Gly His Ala Leu		
405	410	415
Ile Lys Glu Gly Lys Val His Arg Cys Pro Asn Pro Leu Cys Pro Ala		
420	425	430
Lys Arg Phe Glu Ala Ile Arg His Tyr Ala Ser Arg Lys Ala Met Asp		
435	440	445
Ile Gln Gly Leu Gly Glu Lys Leu Ile Glu Lys Leu Leu Glu Lys Gly		
450	455	460
Leu Val Arg Asp Val Ala Asp Leu Tyr Arg Leu Lys Lys Glu Asp Leu		
465	470	475
Val Asn Leu Glu Arg Met Gly Glu Lys Ser Ala Glu Asn Leu Leu Arg		
485	490	495
Gln Ile Glu Glu Ser Lys Gly Arg Gly Leu Glu Arg Leu Leu Tyr Ala		
500	505	510
Leu Gly Leu Pro Gly Val Gly Glu Val Leu Ala Arg Asn Leu Ala Leu		
515	520	525
Arg Phe Gly His Met Asp Arg Leu Leu Glu Ala Gly Leu Glu Asp Leu		
530	535	540
Leu Glu Val Glu Gly Val Gly Glu Leu Thr Ala Arg Ala Ile Leu Asn		
545	550	555
Thr Leu Lys Asp Pro Glu Phe Arg Asp Leu Val Arg Arg Leu Lys Glu		
565	570	575
Ala Gly Val Glu Met Glu Ala Lys Glu Arg Glu Gly Glu Ala Leu Lys		
580	585	590
Gly Leu Thr Phe Val Ile Thr Gly Glu Leu Ser Arg Pro Arg Glu Glu		
595	600	605
Val Lys Ala Leu Leu Arg Arg Leu Gly Ala Lys Val Thr Asp Ser Val		

610

615

620

Ser Arg Lys Thr Ser Phe Leu Val Val Gly Glu Asn Pro Gly Ser Lys
625 630 635 640

Leu Glu Lys Ala Arg Ala Leu Gly Val Pro Thr Leu Ser Glu Glu
645 650 655

Leu Tyr Arg Leu Ile Glu Glu Arg Thr Gly Lys Asp Pro Arg Ala Leu
660 665 670

Thr Ala

<210> 2

<211> 2025

<212> DNA

<213> Thermus sp.

<400> 2

atgaccctag aggaggcccc caggcgctc aacgaactca gggacctgat ccgttaccac 60
aactacctct attacgtctt ggacgcccc gagatctccg acgcccagta cgaccggctc 120
cttagggagc ttaaggagct ggaggagcgc ttcccggc tcaaaagccc cgactcccc 180
acggaacagg tggggcggag gcctctggag gccaccttcc gcccggtgcg ccacccacc 240
cgcatgtact ccctggacaa cgcctttcc ttggacgagg tgagggcctt tgaggagcgc 300
atagagcggg ccctggggcg gaaggggccc ttcccttaca ccgtggagcg caagggtggac 360
ggtcttcgg tgaacctcta ctacgaggag ggcattctcg tctttggggc caccggggc 420
gacggggaga cggggagga ggtgacccag aacctctca ccatccccac cattccccgc 480
cgccctcacgg gcgttccggc cgcctcgag gtccggggcg aggtctacat gcccatacg 540
gccttcctca ggctcaacca ggagctggag gaggcggggg agcgcattttt caaaaacccc 600
aggaacgccc cgcggggc cttgcggcag aaagacccc ggtcacggc caggcggggc 660
ctgagggcca cctttacgc cctggggctg ggcctggagg aaaccgggtt aaaaagccag 720
cacgacccitc tcctatggct aagagagcgg ggcttcccg tggagcacgg ctttaccgg 780
gccctgggg cgggggggt ggaggaggc taccaggct ggctcaagga gaggcggaaag 840
cttcccttg aggccgacgg ggtgggtggc aagctggacg acctcgccct ctggcggggag 900
ctggggtaca cgcggccac ccccgcttc gccctcgccct acaagtccc ggccgaggag 960
aaggagaccc gccttccttc cgtggcccttc caggtggggc ggaccggggc catcacc 1020
gtggcggttc tggagcccgt cttcatagag ggcagcgagg tgagccgggtt caccctccac 1080
aacgagagct tcattgagga gctggacgtg cgcattcgcc actgggtgct ggtccacaag 1140
gcggcggtt tgattcccg agtgctgagg gtccctgaaag agcgcggcac cggggaggag 1200
aagcccatca tctggcccga gaactgcccc gagtgccggcc acgcctcat caaggagggg 1260
aaggtccacc gctgccccaa ccccttgtgc cccgccaagc gcttggaggc catccggcac 1320
tacgcctccc gcaaggccat ggacatccag ggcctggggg agaagctcat agaaaaagctt 1380
ctggaaaagg gcctggccgg ggacgtggcc gaccttacc gcctgaagaa ggaggacctg 1440
gtgaacctgg agcgcattggg ggagaagagc gcagagaacc tccctccca gatagaggag 1500
agcaaggccc gcggccttgg ggcctccctt tacgcctgg gccttccgg ggtgggggag 1560
gtgctggccc ggaacctggc cctccgttc ggcacatgg accgccttct ggaggcgggc 1620

ctcgaggacc tcctggaggt ggagggggtg ggcgagctca ccgcccgggc catcctgaat 1680
accctaaagg acccgagtt ccgggacctg gtgcgcggcc tgaaggaggc cgggggtggag 1740
atggaggcca aagagcggga gggcgaggcc ttgaaggggc tcacccctgt catcaccggg 1800
gagcttcgg gcggccggga ggaggtgaag gcgcctccta ggcggcttgg ggccaagggtg 1860
acggactcgg tgagccgcaa gacgagcttc ctggtggtgg gggagaaccc ggggagcaag 1920
ctggaaaagg cccgcgcctt ggggtcccc accctgagcg aggaggagct ctaccgcctc 1980
attgaggaga ggacgggcaa ggacccaagg gcgcctcacgg cctag 2025

<210> 3
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: probe or primer

<220>
<221> tRNA
<222> (4)
<223> w at position 4 can be T or A

<220>
<221> unsure
<222> (5)
<223> s at position 5 can be C or G

<220>
<221> unsure
<222> (12)
<223> s at position 12 can be C or G

<220>
<221> unsure
<222> (15)
<223> r at position 15 can be G or A

<220>
<221> unsure
<222> (18)
<223> y at position 18 can be T or C

<400> 3
atcwscgacg csgartayga 20

<210> 4

<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: protein

<400> 4
Ile Ser Asp Ala Glu Tyr Asp
1 5

<210> 5
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: probe or primer

<220>
<221> unsure
<222> (3)
<223> s at position 3 can be C or G

<220>
<221> unsure
<222> (6)
<223> s at position 6 can be C or G

<220>
<221> unsure
<222> (8)
<223> k at position 8 can be G or T

<220>
<221> unsure
<222> (9)
<223> s at position 9 can be G or C

<220>
<221> unsure
<222> (12)
<223> s at position 12 can be G or C

<220>
<221> unsure

<222> (15)
<223> y at position 15 can be C or T

<220>
<221> unsure
<222> (18)
<223> r at position 18 can be A or G

<400> 5
ccsgtscksc csacytgraa

20

<210> 6
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: probe or
primer

<220>
<221> unsure
<222> (9)
<223> v at position 9 can be C, G, or A

<220>
<221> unsure
<222> (11)
<223> r at position 11 can be A or G

<220>
<221> unsure
<222> (12)
<223> y at position 12 can be T or C

<220>
<221> unsure
<222> (16)
<223> s at position 16 is C or G

<220>
<221> unsure
<222> (17)
<223> w at position 17 can be A or T

<220>
<221> unsure

<222> (18)
<223> s at position 18 can be G or C

<400> 6
gccttctcva ryttgswscc

20

<210> 7
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: probe or
primer

<400> 7
Phe Gln Val Gly Arg Thr Gly
1 5

<210> 8
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: probe or
primer

<400> 8
Gly Ser Lys Leu Glu Lys Ala
1 5

<210> 9
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: probe or
primer

<400> 9
gcgatttcat atgaccctag aggaggccccg

30

<210> 10
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: probe or primer

<400> 10
gcgggatccg aggcccttgg aaaaaaaaaaaaaaaa 29

<210> 11
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: probe or primer

<400> 11
aaaaccctgt tccagcgtct gcgggtttgc gtc 33

<210> 12
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: probe or primer

<400> 12
agttgtcata gtttgatcct ctagtctggg 30

<210> 13
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: probe or primer

<400> 13
ccctgttcca gcgtctgcgg tgttgcgtt

29

<210> 14
<211> 59
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: probe or primer

<400> 14
gggacaaggc cgcaacgcgtt caacagtatc aaactaggag atcagaccc 59

<210> 15
<211> 184
<212> PRT
<213> Thermus aquaticus

<220>
<221> UNSURE
<222> (18)..(120)
<223> Xaa at positions 18-120 is any amino acid

<220>
<221> UNSURE
<222> (126)..(172)
<223> Xaa at positions 126-172 is any amino acid

<400> 15
Tyr Thr Val Glu His Lys Val Asp Gly Leu Ser Val Asn Leu Tyr Tyr
1 5 10 15

Glu Xaa
20 25 30

Xaa
35 40 45

Xaa
50 55 60

Xaa
65 70 75 80

Xaa
85 90 95

Xaa
100 105 110

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Glu Glu Thr Gly Xaa Xaa Xaa
115 120 125

Xaa
130 135 140

Xaa
145 150 155 160

Xaa Pro Phe Glu Ala
165 170 175

Asp Gly Val Val Val Lys Leu Asp
180

<210> 16

<211> 187

<212> PRT

<213> Thermus flavus

<220>

<221> UNSURE

<222> (18)..(120)

<223> Xaa at positions 18-120 is any amino acid

<220>

<221> UNSURE

<222> (129)..(175)

<223> Xaa at positions 129-175 is any amino acid

<400> 16

Tyr Thr Val Glu His Lys Val Asp Gly Leu Ser Val Asn Leu Tyr Tyr
1 5 10 15

Glu Xaa
20 25 30

Xaa
35 40 45

Xaa Xaa

50

55

60

Xaa
65 70 75 80

Xaa
85 90 95

Xaa
100 105 110

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Glu Glu Val Glu Arg Glu Gly
115 120 125

Xaa
130 135 140

Xaa
145 150 155 160

Xaa Pro
165 170 175

Phe Glu Ala Asp Gly Val Val Val Lys Leu Asp
180 185

<210> 17

<211> 184

<212> PRT

<213> *Thermus filiformis*

<220>

<221> UNSURE

<222> (18)..(120)

<223> Xaa at positions 18-120 is any amino acid

<220>

<221> UNSURE

<222> (126)..(172)

<223> Xaa at positions 126-172 is any amino acid

<400> 17

Tyr Thr Val Glu His Lys Val Asp Gly Leu Ser Val Asn Leu Tyr Tyr
1 5 10 15

Glu Xaa
20 25 30

Xaa			
35	40	45	
Xaa			
50	55	60	
Xaa			
65	70	75	80
Xaa			
85	90	95	
Xaa			
100	105	110	
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Glu Glu Ser Gly Xaa Xaa Xaa			
115	120	125	
Xaa			
130	135	140	
Xaa			
145	150	155	160
Xaa Pro Phe Glu Ala			
165	170	175	
Asp Gly Val Val Val Lys Met Asp			
180			

<210> 18
<211> 184
<212> PRT
<213> *Thermus filiformis*

<220>
<221> UNSURE
<222> (18)..(120)
<223> Xaa at positions 18-120 is any amino acid

<220>
<221> UNSURE
<222> (126)..(172)
<223> Xaa at positions 126-172 is any amino acid

<400> 18

Tyr	Thr	Val	Glu	His	Lys	Val	Asp	Gly	Leu	Ser	Val	Asn	Leu	Tyr	Tyr				
1									10					15					
Glu Xaa																			
					20						25						30		
Xaa																			
					35						40						45		
Xaa																			
					50						55						60		
Xaa																			
					65						70						75	80	
Xaa																			
					85						90						95		
Xaa																			
					100						105						110		
Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Glu Glu Ser Gly Xaa Xaa Xaa																			
					115						120						125		
Xaa																			
					130						135						140		
Xaa																			
					145						150						155	160	
Xaa Pro Phe Glu Ala																			
					165						170						175		
Asp Gly Val Val Val Lys Leu Asp																			
					180														

<210> 19
<211> 184
<212> PRT
<213> *Thermus sp.*

<220>
<221> UNSURE
<222> (18)..(120)
<223> Xaa at positions 18-120 is any amino acid

<220>

<221> UNSURE
<222> (126)..(172)
<223> Xaa at positions 126-172 is any amino acid

<400> 19

Tyr Thr Val Glu His Lys Val Asp Gly Leu Ser Val Asn Leu Tyr Tyr
1 5 10 15

Glu Xaa
20 25 30

Xaa
35 40 45

Xaa
50 55 60

Xaa
65 70 75 80

Xaa
85 90 95

Xaa
100 105 110

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Glu Glu Ser Gly Xaa Xaa Xaa
115 120 125

Xaa
130 135 140

Xaa
145 150 155 160

Xaa Pro Phe Glu Ala
165 170 175

Asp Gly Val Val Val Lys Leu Asp
180

<210> 20

<211> 184

<212> PRT

<213> Thermus sp.

<220>

<221> UNSURE

<222> (18) .. (120)

<223> Xaa at positions 18-120 is any amino acid

<220>

<221> UNSURE

<222> (126) . . (172)

<223> Xaa at positions 126-172 is any amino acid

<400> 20

Tyr Thr Val Glu His Lys Val Asp Gly Leu Ser Val Asn Leu Tyr Tyr
 1 5 10 15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Glu Glu Ser Gly Xaa Xaa Xaa
 115 120 125

Xaa Pro Phe Glu Ala

Asp Gly Val Val Val Lys Leu Asp

<210> 21

<211> 184
<212> PRT
<213> Thermus sp.

<220>
<221> UNSURE
<222> (18)..(120)
<223> Xaa at positions 18-120 is any amino acid

<220>
<221> UNSURE
<222> (126)..(172)
<223> Xaa at positions 126-172 is any amino acid

<400> 21
Tyr Thr Val Glu His Lys Val Asp Gly Leu Ser Val Asn Leu Tyr Tyr
1 5 10 15

Glu Xaa
20 25 30

Xaa
35 40 45

Xaa
50 55 60

Xaa
65 70 75 80

Xaa
85 90 95

Xaa
100 105 110

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Glu Glu Ser Gly Xaa Xaa Xaa
115 120 125

Xaa
130 135 140

Xaa
145 150 155 160

Xaa Pro Phe Glu Ala
165 170 175

Asp Gly Val Val Val Lys Leu Asp

180

<210> 22

<211> 184

<212> PRT

<213> Thermus aquaticus

<220>

<221> UNSURE

<222> (18)..(120)

<223> Xaa at positions 18-120 is any amino acid

<220>

<221> UNSURE

<222> (126)..(172)

<223> Xaa at positions 126-172 is any amino acid

<400> 22

Tyr Thr Val Glu Arg Lys Val Asp Gly Leu Ser Val Asn Leu Tyr Tyr
1 5 10 15

Glu Xaa
20 25 30

Xaa
35 40 45

Xaa
50 55 60

Xaa
65 70 75 80

Xaa
85 90 95

Xaa
100 105 110

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Glu Glu Thr Gly Xaa Xaa Xaa
115 120 125

Xaa
130 135 140

Xaa Xaa

145

150

155

160

Xaa Pro Phe Glu Ala
165 170 175

Asp Gly Val Val Val Lys Leu Asp
180

<210> 23

<211> 187

<212> PRT

<213> Thermus flavus

<220>

<221> UNSURE

<222> (18)..(120)

<223> Xaa at positions 18-20 is any amino acid

<220>

<221> UNSURE

<222> (129)..(175)

<223> Xaa at positions 129-175 is any amino acid

<400> 23

Tyr Thr Val Glu His Lys Val Asp Gly Leu Ser Val Asn Leu Tyr Tyr
1 5 10 15

Glu Xaa
20 25 30

Xaa
35 40 45

Xaa
50 55 60

Xaa
65 70 75 80

Xaa
85 90 95

Xaa
100 105 110

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Glu Glu Val Glu Arg Glu Gly
115 120 125

Xaa
130 135 140

Xaa
145 150 155 160

Xaa Pro
165 170 175

Phe Glu Ala Asp Gly Val Val Val Lys Leu Asp
180 185

<210> 24

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Peptide

<220>

<221> VARIANT

<222> (2)

<223> X at position 2 is any amino acid

<400> 24

Lys Xaa Asp Gly

1